To conclude, Zandvakili and Kohn (2015) present the first definitive evidence that spiking coordination does indeed have a modest impact on the efficacy of transmission in the neocortex. They also show that the effects of spiking coordination in a source population do not propagate past the input layers of the target population, at least in the early visual cortex. However, it remains possible that more subtle aspects of coding beyond the input layers are affected. Importantly, they have established an empirical framework in which such hypotheses can be directly tested at the level where cortical computations are performed: spiking activity in neuronal populations.

REFERENCES

Abeles, M. (1991). Corticonics: Neural circuits of the cerebral cortex (Cambridge University Press).

Alonso, J.M., Usrey, W.M., and Reid, R.C. (1996). Nature *383*, 815–819.

Bruno, R.M., and Sakmann, B. (2006). Science *312*, 1622–1627.

Einevoll, G.T., Kayser, C., Logothetis, N.K., and Panzeri, S. (2013). Nat. Rev. Neurosci. *14*, 770–785.

Fries, P. (2005). Trends Cogn. Sci. *9*, 474–480.

Fries, P. (2009). Annu. Rev. Neurosci. *32*, 209–224.

Gray, C.M. (1999). Neuron *24*, 31–47, 111–125.

Jia, X., Tanabe, S., and Kohn, A. (2013). Neuron *77*, 762–774.

Kojima, S., and Doupe, A.J. (2009). J. Neurosci. *29*, 4782–4793.

Kumar, A., Rotter, S., and Aertsen, A. (2010). Nat. Rev. Neurosci. *11*, 615–627.

Renart, A., de la Rocha, J., Bartho, P., Hollender, L., Parga, N., Reyes, A., and Harris, K.D. (2010). Science *327*, 587–590.

Roy, S.A., and Alloway, K.D. (2001). J. Neurosci. *21*, 2462–2473.

Roy, S.A., Dear, S.P., and Alloway, K.D. (2001). J. Neurosci. *21*, 1795–1808.

Shadlen, M.N., and Newsome, W.T. (1998). J. Neurosci. *18*, 3870–3896.

Zandvakili, A., and Kohn, A. (2015). Neuron *87*, this issue, 827–839.

It All Depends on the Context, but Also on the Amygdala

Yarden Cohen[1] and Rony Paz[1,*]
[1]Department of Neurobiology, Weizmann Institute of Science, Rehovot 76100, Israel
*Correspondence: rony.paz@weizmann.ac.il
http://dx.doi.org/10.1016/j.neuron.2015.08.012

Behavioral flexibility requires the brain to maintain and rely on cognitive contexts for dictating appropriate responses. Saez et al. (2015) demonstrate that such abstract rule-based representations co-exist in prefrontal cortices and in the amygdala, with the latter being surprisingly crucial for correct performance.

The hallmark of successful learning is the ability to produce appropriate behavioral action in response to a specific external stimulus that has an emotional value. However, the same stimulus can produce different, and in times even opposite, actions and emotional responses—depending on the exact context in which it is encountered. For example, an approaching lion induces an approach response accompanied by excitement and elevated mood when encountered in the zoo. If encountered in the wild, however, the exact same lion stimulus would induce fear and pose an immediate threat that calls for a dramatically different action—a flight response. Sometimes contexts can even be un-cued and instead simply be defined by a cognitive process. For example, the same hand of cards can have a different meaning depending upon one's understanding of the rules of the game being played.

This ability to react in accordance with multiple particular contexts requires the adoption of a complete, context-specific behavioral set. Such behavioral sets allow rapid switching between different behaviors and actions, depending on the context in which the stimulus is experienced. Interestingly, these commonly used adaptive behavioral sets are an antipode of classical conditioning theories. In these learning schemes, the internal value assigned to a conditioned stimulus (CS) is altered when the stimulus is coupled to a meaningful (appetitive or aversive) unconditioned outcome (US). The conditioned response (CR) is the external manifestation of the internal value reflecting the learned association, and it is limited to the specific stimulus. Classical theoretical models of conditioning (Rescorla and Wagner, 1972; Sutton and Barto, 1998) predict that a change in the reinforcement value would necessitate re-learning of the new association. Conversely, context-based adaptation means that not only the response to a single stimulus is changed without having to re-learn the CS-US pairing again and again, but that responses to other stimuli also change simultaneously—reflecting the switch to a wholly different behavioral set. Acquisition and maintenance of such complete behavioral sets offer a clear evolutionary advantage. Without a representation of behavioral sets, an organism would have to re-learn all CS-US associations in every context again and again,

a highly energetic and time-consuming effort, posing a serious survival threat.

Context-based switches can also greatly assist in regulation of emotion. We need not gradually learn to feel safe next to each type of, otherwise dangerous, animal in the zoo. Rather, switching to the safe context provides a regulatory mechanism of emotional control. Clinical implications of dysfunction in such a mechanism may include uncontrolled emotional swings due to the need to re-learn associations repeatedly in the absence of appropriate behavioral sets to switch to. Conversely, ossified clinging to one irrelevant context can result in a host of maladaptive behaviors and emotional inappropriate responses. Such can be the case, for example, in mood disorders like depression and post-traumatic stress.

Traditionally, prefrontal cortices (PFCs) were considered to acquire and hold information about rules (Buckley et al., 2009; Wallis et al., 2001; Wilson et al., 2014). These textbook findings make intuitive sense when considering the development and expansion of prefrontal regions in primates and their advanced ability to use abstract rules and smartly switch between behavioral sets according to the situation. The amygdala, on the other hand, is traditionally associated with facilitating and encoding the associations between a stimulus and its outcome; in particular, when this outcome has an emotional value, and even more specifically when the value is a negative one. This "fear-centered" view of the amygdala has dominated the field and oriented amygdala research, even if not explicitly. Yet a recent surge of electrophysiological studies in primates has revealed a much more complex picture, showing, for example, that primate amygdala neurons hold information about both positive and negative values (Paton et al., 2006), signal safety (Genud-Gabai et al., 2013), and expectation (Belova et al., 2007); process spatial cues to orient attention (Peck et al., 2013); compute prediction errors (Klavir et al., 2013); and signal internally generated reward-based choices (Hernádi et al., 2015), as well as mediate generalization of learning (Resnik and Paz, 2015) and interact closely with the prefrontal cortex to guide statistical learning (Livneh and Paz, 2012).

Nevertheless, taking active part in maintaining abstract information about context and rules, and hence allowing the use of behavioral sets to guide action, remained the domain of PFCs alone.

Considering the evolutionary expansion and extensive bidirectional connectivity between the amygdala and PFCs in primates, Saez et al. (2015) provide data that challenge this view. To do so, they taught monkeys to collect fluid reward in two CS-US pairing sets, or contexts. In each context, the reward was administered exclusively after one of the two stimuli. The reinforcement contingencies of the two CSs reversed back and forth many times in experiments to create the two contexts. Sometimes, a clear additional visual cue marked the context within a trial, but on the majority of trials, context was un-cued, so monkeys had to represent the context as an internal cognitive variable. Monkeys indeed switched their behavior reflecting reward expectancy for the two CSs after every block switch. Remarkably, once the monkeys had experienced that one CS had switched its reinforcement contingency, they inferred correctly that the other CS had switched its contingency after the block switch, even though they had not yet experienced that CS as having switched. Indeed, monkeys' reward expectancy behavior reflected an abrupt and persistent transition between contexts. This was the case even for the first CS-US reversal experienced in an experiment. Monkeys thereby exhibited clear cognitive, internally represented, context usage. In simple terms, they understood the abstract rule and task structure and did not need to learn de novo the associations per each new stimulus (as expected from classical learning theories).

To probe the differential representations of abstract context, stimulus identity, and reward expectancy, Saez et al. (2015) recorded single units from the amygdala and two PFC regions: the anterior cingulate (ACC) and the orbitofrontal cortex (OFC), both densely interconnected with the amygdala (Ghashghaei et al., 2007). They found that stimulus identity and reward expectancy were represented extensively in all three regions, and they identified for the first time that information about context is represented in ACC and OFC.

Yet to their (and our) surprise, information about context was encoded by amygdala neurons as well and, moreover, to a similar extent as in the PFC. Importantly, this information was available even before stimulus onset, as can be expected from knowledge about abstract context. To reach these conclusions, Saez et al. (2015) used both the standard approach of quantifying unit responses in relation to task parameters, but they also exploited the use of linear decoders. This approach has several benefits: first, it allows aggregating information from all recorded neurons and hence probing population coding within a region; second, even if we do not know how the neurons precisely code information (a fundamental problem in all of neuroscience), the decoder approach circumvents this elegantly and probes the available information; third, and importantly, it allows training the decoder on a set of the trials to distinguish between two conditions and then testing whether this trained decoder can distinguish between two conditions on held-out trials. One can then draw conclusions about similarities and discrepancies in the amount of information across conditions, and thus reach further insights into what the neurons actually code for.

For example, if the decoder was trained on trials without an explicit contextual cue but was tested on trials with it, the decoding performance for context was just as good. This is strong evidence that the information held is indeed about abstract context. In a similar way, it was shown that decoding context does not rely on recent memory trace of specific stimuli. Instead, and importantly, the context representation reflects the linked sets of CS-US associations that defined each context—the relevant "behavioral set" for this task. This is a powerful demonstration for coding of abstract rules rather than traditional reinforcement-driven computations only, which is especially surprising to observe in the amygdala.

Luckily and naturally, monkeys also performed errors. An intelligent observation of how decoding properties signal information during such trials allowed the authors to determine that errors could not be accounted for by failures in stimulus identification encoding. Rather, errors were correlated with a decrease

in context information, which was seen most prominently in the amygdala. Hence, not only is context represented in the amygdala, but failure to maintain this information might contribute to behavioral failures. It reinforces the interpretation that the abstract context signals observed in the amygdala are more than mere correlates projected from previously identified cortices, but that the amygdala takes active part (causal?) in contributing to make the final decision and computing action and outcome expectancies.

Such findings might have important clinical implications. If the amygdala holds abstract context and behavioral sets that are responsible for appropriate action (yet at the same time is highly modulated by emotional reinforcement, as we know from many studies), then minor failures in the coding process can result in psychopathological conditions. From a computational perspective, it requires careful separation of transient reinforcement-driven computations from abstract rule-based representations and their correct integration when required—all within the same network.

These findings strongly suggest, for the first time, that the amygdala actively participates in maintenance of abstract cognitively relevant information. The next steps therefore should be to identify the mechanisms and brain regions by which learning of the context, or rule abstraction, is mediated. For example, does the amygdala participate in learning the context abstraction as well? Alternatively, behavioral sets can be learned via prefrontal circuits as commonly hypothesized and only later projected into the amygdala, allowing this information to interact with transient reinforcement properties and stimulus-outcome associations to reach the correct action in each specific real-world experience.

These challenges require three major steps in our view. The first is the simultaneous recordings and analyses of activity in multiple brain regions (and multiple neurons within them). This would allow identifying the information not only within a region, but also the information transferred across regions and potentially even maintained across regions. It is technologically plausible, but the application and development of analyses and analytical methods is still ongoing.

The second is the use of direct analyses approaches that go beyond decoding and/or the use and interpretation of several decoder options. The elegant use of decoding in this study allowed Saez et al. (2015) to ask and answer several important aspects about the information maintained, but decoders do not directly address the question of how exactly information is coded, and how exactly downstream networks read it. Extending to different types of decoders and directly observing information in precise spiking properties might offer more insights. We note that this is a general major challenge in the field.

Finally, electrophysiological studies of complex rule switching, especially during learning of new rules and contexts, pose a major challenge for neuroscientists, mainly due to variability and erratic behavior that imposes major confounds on interpretations of neural findings. Moreover, realistic rule-based scenarios can entail much larger stimulus-outcome sets, leading to more complex and harder-to-control experimental conditions. Following this important study by Saez et al. (2015), the next big challenge would be to attempt to study and track the brain substrates of learning, executing and switching between complex rules, and identifying large circuit-based interactions.

REFERENCES

Belova, M.A., Paton, J.J., Morrison, S.E., and Salzman, C.D. (2007). Neuron *55*, 970–984.

Buckley, M.J., Mansouri, F.A., Hoda, H., Mahboubi, M., Browning, P.G., Kwok, S.C., Phillips, A., and Tanaka, K. (2009). Science *325*, 52–58.

Genud-Gabai, R., Klavir, O., and Paz, R. (2013). J. Neurosci. *33*, 17986–17994.

Ghashghaei, H.T., Hilgetag, C.C., and Barbas, H. (2007). Neuroimage *34*, 905–923.

Hernádi, I., Grabenhorst, F., and Schultz, W. (2015). Nat. Neurosci. *18*, 461–469.

Klavir, O., Genud-Gabai, R., and Paz, R. (2013). Neuron *80*, 1290–1300.

Livneh, U., and Paz, R. (2012). Neuron *75*, 133–142.

Paton, J.J., Belova, M.A., Morrison, S.E., and Salzman, C.D. (2006). Nature *439*, 865–870.

Peck, C.J., Lau, B., and Salzman, C.D. (2013). Nat. Neurosci. *16*, 340–348.

Rescorla, R.A., and Wagner, A.R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In Classical conditioning, A.H. Black and W.F. Prokasy, eds. (Appleton-Century-Crofts), pp. 64–99.

Resnik, J., and Paz, R. (2015). Nat. Neurosci. *18*, 188–190.

Saez, A., Rigotti, M., Ostojic, S., Fusi, S., and Salzman, C.D. (2015). Neuron *87*, this issue, 869–881.

Sutton, R.S., and Barto, A.G. (1998). Reinforcement Learning: An Introduction (MIT press).

Wallis, J.D., Anderson, K.C., and Miller, E.K. (2001). Nature *411*, 953–956.

Wilson, R.C., Takahashi, Y.K., Schoenbaum, G., and Niv, Y. (2014). Neuron *81*, 267–279.